Andrew Schechtman-Rook

rook166@gmail.com   https://github.com/AndrewRook
(917)-836-4267

TECHNICAL SKILLS

Programming: Python (numpy, scipy, pandas, sklearn, xgboost, matplotlib), C++, shell scripting
Model Development & Deployment: GBMs, random forests, linear & logistic regression, nonlinear optimization, parallel & distributed computing, containerization, APIs
Databases, Orchestration, Web Design: Flask, MySQL/PostgreSQL, Prefect, Snowflake
Cloud Computing/Devops: AWS, CircleCI, Jenkins, GitHub Actions

WORK EXPERIENCE

Capital One   2014-Present
Principal Associate Data Scientist— Staff Data Scientist/Machine Learning Engineer

- Built a RAG app using Mixtral and ChromaDB to generate business-specific SQL queries from plaintext prompts, achieving accuracies of over 80% of a test dataset of real queries.
- Prototyped a from-scratch transformer model using Pytorch trained on historical analyst SQL queries.
- Deployed the first cloud-based credit card underwriting model in the company via a dockerized Python API, with an estimated incremental value of 35 million dollars per year.
- Created the longest-lived, most successful internal data science tool in the company, used in production models by dozens of data scientists across multiple lines of business.
- Programmed and deployed an interactive course completion dashboard using Flask and dc.js to provide progress reports to individual students as well as company leadership.

University of Wisconsin-Madison   2007-2014
Graduate Student — Postdoctoral Researcher

- Created a fast Voronoi Tessellation algorithm in Python to adaptively bin images, preserving spatial resolution while maximizing signal in images with over one million pixels.
- Developed non-linear Levenberg-Marquardt χ^2 fitting algorithms using a combination of Python and C++ to constrain models of spiral galaxies to data.
- Employed on-campus distributed computing resources to perform large-scale modeling in parallel, using over 20 years of computer time in 1 month.
- Assembled a hybrid C++/Python processing pipeline to clean, register, and mosaic thousands of high-resolution images with minimal user intervention, resulting in a factor of 10+ increase in analysis precision.

EDUCATION

PhD, Astronomy, University of Wisconsin-Madison	December 2013
MS, Astronomy, University of Wisconsin-Madison	June 2009
BS, Astronomy, Case Western Reserve University	May 2007